



```

ss8440c cli% showldch -lformat set -linfo pdpos,pdch,pdid vv_test.usr.0
row set      ch0      ch1      ch2      ch3      ch4      ch5      ch6      ch7
0 0 3:2:0-93-28 0:2:0-118-2 2:2:0-117-22 1:2:0-117-14 3:4:0-89-30 1:4:0-116-16 2:0:0-116-20 0:0:0-0-0
1 0 3:2:0-94-28 0:2:0-119-2 2:2:0-118-22 1:2:0-118-14 3:4:0-90-30 1:4:0-117-16 2:0:0-117-20 0:0:0-114-0
2 0 3:6:0-91-32 1:0:0-114-12 0:4:0-116-4 2:4:0-117-24 3:0:0-92-26 0:0:0-115-0 2:0:0-118-20 1:4:0-118-16
ss8440c cli% showldch -lformat set -linfo pdpos,pdch,pdid vv_test.usr.1
row set      ch0      ch1      ch2      ch3      ch4      ch5      ch6      ch7
0 0 3:5:0-111-31 1:5:0-112-17 2:5:0-112-25 0:5:0-111-5 1:1:0-112-13 2:1:0-112-21 0:1:0-112-1 3:1:0-112-27
1 0 3:5:0-112-31 1:5:0-113-17 2:5:0-113-25 0:5:0-112-5 1:1:0-113-13 2:1:0-113-21 0:1:0-113-1 3:1:0-113-27
2 0 3:5:0-113-31 1:5:0-114-17 2:5:0-114-25 0:5:0-113-5 1:1:0-114-13 2:1:0-114-21 0:1:0-114-1 3:1:0-114-27
ss8440c cli% showldch -lformat set -linfo pdpos,pdch,pdid vv_test.usr.2
row set      ch0      ch1      ch2      ch3      ch4      ch5      ch6      ch7
0 0 4:0:0-108-33 6:0:0-110-47 5:0:0-110-41 7:0:0-164-55 6:2:0-115-49 4:2:0-114-35 5:2:0-115-43 7:2:0-166-57
1 0 4:0:0-112-33 6:0:0-114-47 5:0:0-114-41 7:0:0-165-55 6:2:0-116-49 4:2:0-115-35 5:2:0-116-43 7:2:0-167-57
2 0 4:0:0-113-33 6:0:0-115-47 5:0:0-115-41 7:0:0-166-55 6:4:0-113-51 4:4:0-116-37 5:4:0-114-45 7:2:0-168-57
ss8440c cli% showldch -lformat set -linfo pdpos,pdch,pdid vv_test.usr.3
row set      ch0      ch1      ch2      ch3      ch4      ch5      ch6      ch7
0 0 7:3:0-109-50 6:3:0-109-50 5:3:0-110-44 4:3:0-109-30 7:1:0-113-56 4:1:0-113-40 6:5:0-113-52 5:5:0-112-46
1 0 7:3:0-112-58 6:3:0-113-50 5:3:0-113-44 4:3:0-112-36 7:1:0-114-56 4:1:0-114-40 6:5:0-114-52 5:5:0-113-46
2 0 7:3:0-113-58 4:3:0-113-36 6:3:0-114-50 5:3:0-114-44 5:1:0-116-42 6:1:0-116-48 4:5:0-116-38 7:5:0-117-59

```

分析:

由vv\_test.usr.0作为例子进行分析

对于第一个set, 选择了28号, 2号, 22号, 14号, 30号, 16号, 20号, 0号磁盘 (PDID) 组成了一个raid,

其中2号以及0号磁盘位于cage0;

14号以及16号磁盘位于cage 1;

28号以及30号磁盘位于cage 3;

20号以及22号磁盘位于cage 2;

所有的PDID都是偶数盘, 即node 0选择node pair对连接的盘柜中的8块偶数盘组成了一个raid,此时当一个cage掉电, 比如cage 3掉电后, 28号以及30号磁盘此时无法使用, 因为做的是raid 6,允许坏两块盘, 因此数据还是能够正常访问的, 但是如果坏了两个cage, raid结构损坏, 数据无法访问;

注意:

- LD由单个node 管理, 当一个node 损坏以后, 另一个node会进行接管, 如vv\_test.usr.0由node 0接管, node 0 down了以后, 会由1,2, 3进行接管
- 奇数node只能管理奇数盘, 偶数node 只能管理偶数盘
- 对于raid 6的写惩罚: 写入一个新的数据, raid1的写惩罚为2, raid 5的写惩罚为4, raid6的写惩罚为6.66, (当raid set长度为8的时候, 有三分之二的的数据需要更新2个parity (此时为写惩罚为6: 读取旧数据, 两个奇偶校验, 写入旧数据, 写入两个奇偶校验) 而数据块的1/3用于计算3个奇偶校验 (此时写惩罚为8, 读取旧数据, 三个奇偶校验, 写入新数据与3个奇偶校验) 即 $2/3*6+1/3*8=6.66$ )
- raid 5的写惩罚: 随机写入RAID 5卷将需要四个后端I/O。该过程开始于读取旧数据块和奇偶校验块 (两次后端读取), 并通过对新数据, 旧数据和旧奇偶校验值进行XOR运算来计算新奇偶校验信息。然后将新数据和新奇偶校验写入后端, 从而创建两个后端写入。
- raid 6的空间使用率: 比如本实验:  $6/8=75%$

```

ss8440c cli% showvv -s vv_test
-----Snp-----Usr-----Total-----
--(MiB)-- --(% VSize)-- --(MiB)-- --(% VSize)-- --(MiB)-- --Efficiency--
Id Name  Prov Compr Dedup Type Rsvd Used Wrn Lim Rsvd Used Wrn Lim Rsvd Used HostWr VSize Compact Compress
-----
1 total  0 0 0 0 51200 51200 0 0 51200 51200 0 0 51200
ss8440c cli% showvv -r vv_test
-----Snp-----Total-----
--(MiB)-- --(% VSize)-- --(MiB)-- --(% VSize)-- --(MiB)-- --Efficiency--
Id Name  Prov Compr Dedup Type RawRsvd Rsvd RawRsvd Rsvd VSize
-----
282 vv_test full NA NA base 0 0 68267 51200 68267 51200 51200
1 total  0 0 68267 51200 68267 51200 51200

```

6、LD只能使用同一个node pair对下的cage下的磁盘进行创建, 即3PAR的LD 不能跨node pair对;